

turning knowledge into practice

Web Based Metadata Tracking System Designed for the National Survey on Drug Use and Health

Nanthini Ganapathi, Susan K. Myers, Inga B. Allred

March 14, 2006



RTI International is a trade name of Research Triangle Institute

3040 Cornwallis Road
Phone 919-541-7143

■ P.O. Box 12194

■ Research Triangle Park, North Carolina, USA 27709
Fax 919-541-6178

■ e-mail nanthini@rti.org

NSDUH Project

- National Survey on Drug Use and Health (**NSDUH**) – formerly National Household Survey on Drug Abuse (**NHSDA**). Sponsored by SAMHSA
- Primary source of information on the prevalence, patterns, and consequences of alcohol, tobacco, and illegal drug use & abuse in the general U.S. civilian non institutionalized population, age **12** and older
- Data are used by policymakers and researchers to assess the prevalence and correlates of licit and illicit drug use, to identify and monitor trends in substance use, and analyze differences by population subgroups
- Screen approx **170,000** households annually, perform approx **67,500** interviews annually
- **50** states + **DC** survey, year round survey.
- Approx **750** field interviewers, geographically dispersed, urban and rural

Needs

- Large scale survey
- Spans for many years
- Lots of effort, time and money spent on collecting this data
- Analysts need to understand the data and what it represents
- Humanly impossible to keep track of these variables and their documentation, manually

Goals

- To keep the data documentation available at all times to the analysts
- To keep the data description current and consistent
- To make the analysts responsible for their own variable documentation
- To fully automate sharing of data including locating, retrieving and exchanging the data
- To automate the creation of electronic and paper code books

Variable Tracking System

Variable Tracking System (VTS) was designed to manage metadata

- About 6000 variables are created and documented each study year.
- 60 staff create and document these variables
- Three formal codebook types are delivered to SAMHSA each study year
 - **Raw**
 - **Analytic**
 - **Public Use File (PUF)**

ISO Metadata General Principles

- Framework
- Classification
- Registry Metamodel and Basic Attributes
- Formulation of Data Definitions
- Naming and Identification Principles
- Registration
- http://metadata-standards.org/11179-1/ISO-IEC_11179-1_1999_IS_E.pdf

Codebook Elements

- Section heading
- Variable Name
- Comment or note about the variable
- SAS label
- Frequency levels (distinct, range or combination of each)
- Frequency level definitions
- Associated footnote

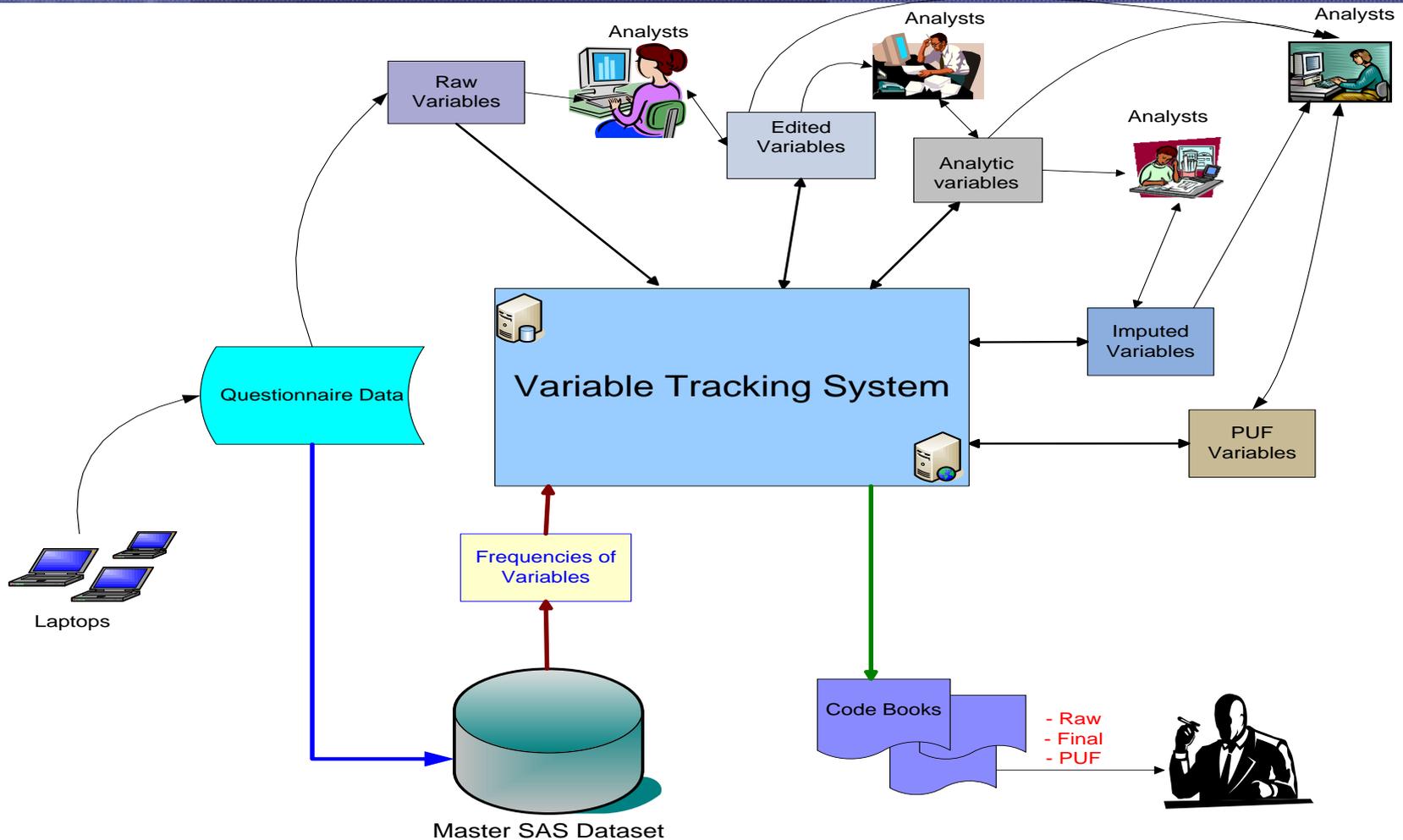
CodeBook Sample

NON-CORE DEMOGRAPHICS

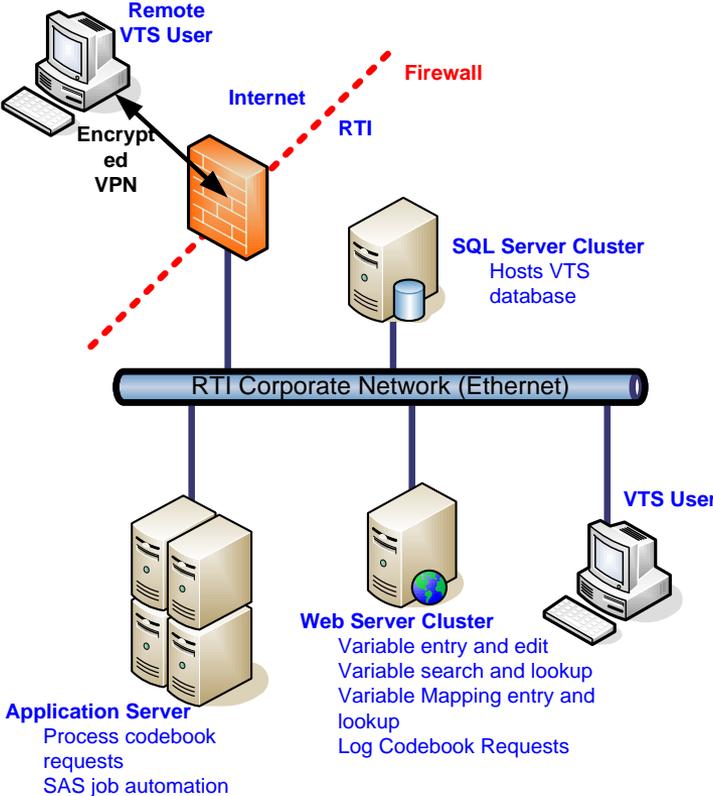
PAGE: 603

LABEL	LEN	DESCRIPTION	FREQ	%
----	---	-----	----	-
(QD17)		The next questions are about school. Are you now attending or are you currently enrolled in school? By "school," we mean an elementary school, a junior high or middle school, a high school, or a college or university. Please include home schooling as well.		
SCHENRL	2	NOW ENROLLED IN ANY SCHOOL		
		1 = Yes.....	26263	48.56
		2 = No.....	27765	51.34
		5 = Yes LOGICALLY ASSIGNED (LFSCHWH2=601).....	12	0.02
		11 = Yes (SCHDSKIP = 30).....	8	0.01
		94 = DON'T KNOW.....	4	0.01
		97 = REFUSED.....	4	0.01
		98 = BLANK (NO ANSWER).....	23	0.04

Life Cycle of a Variable



Technical Aspects



Key Features of VTS

- Data Documentation Entry / Edit /Search
- Data Documentation Tracking
- Variable Delivery / History
- Interrelationship or Mapping of Variables
- Automated CodeBook Creation

VTS Main User Interface

NSDUH Data Processing

NSDUH Variable Tracking

[Variable Lookup](#)
[Map Info Lookup](#) **NEW!**
[Documentation Edit History](#)
[Documentation Edit History \(Levels\)](#)

NSDUH Variable Entry

[Variable Documentation Entry](#)
[Import Documentation](#)
[Codebook Request](#)
[Enter Map Info for Variables](#) **NEW!**

NSDUH Variable Delivery

[Master File Variable Delivery](#)
[Master File Variable Delivery History](#)

NSDUH Deliverables

[NSDUH Deliverables Entry](#)
[NSDUH Deliverables Look up](#)

Special Functions

[Change Module Names](#)
[Module Order](#)
[Order Variables](#)
[Move Variables to a New Section](#)

Project Schedule

[Data Processing Schedule \(Microsoft Project\)](#)
[Data Editing Schedule](#)
[Upload Schedule](#)

Reference

[Reference Library Look up](#)
[Reference Library Entry](#)
[Contacts Information](#)
[STRD Reference Manual](#) [Upload Manual](#)
[Data Processing Website Documentation](#)

Event Codes

[Final Screening and Interview codes \(detailed\)](#)
[Final Screening and Interview codes](#)
[Control system Event Code Lookup](#)

Reports

[Person Selection Report](#)

Lookups

[Remarks Lookup](#)
[Trends Look up](#)

Entering Metadata

Add New Variable (Label)	
Variable Name	<input type="text"/> (No Spaces)
File Type	Raw: <input type="checkbox"/> Analytic: <input type="checkbox"/> PUF: <input type="checkbox"/> Supplemental: <input type="checkbox"/>
Raw Section	Select One <small>For New Section choose New section from this list.</small>
Raw Title Indicator	Does this variable start a new section in the Raw Codebook? Yes <input type="radio"/> No <input checked="" type="radio"/>
Analytic Section	Select One <small>For New Section choose New section from this list.</small>
Analytic Title Indicator	Does this variable start a new section in the Analytic Codebook? Yes <input type="radio"/> No <input checked="" type="radio"/>
PUF Section	Select One <small>For New Section choose New section from this list.</small>
PUF Title Indicator	Does this variable start a new section in the PUF Codebook? Yes <input type="radio"/> No <input checked="" type="radio"/>
Supplemental Section	Select One <small>For New Section choose New section from this list.</small>

Footnote	Select One
Variable Derived From	<input type="text"/> (Related Variable)
Category	1. Defined Levels <input type="radio"/> 2. No Freq/No Levels <input type="radio"/> 3. Continuous/Range <input type="radio"/>
Type	Numeric <input type="radio"/> Alpha <input type="radio"/>
Year	2005
Length	<input type="text"/> * This is NOT SAS Length. This is the maximum character width of the value.
Label	<input type="text"/>
Low Range	<input type="text"/>
Hi Range	<input type="text"/>
Question (comment)	<div style="border: 1px solid gray; height: 80px; width: 100%;"></div> <small>Please format the comments the way you would like to see in the code book including empty lines (60 chr max per line). Use TextPad/Text Editor. Pl. Do not use Word/Word Perfect.</small>
<input type="button" value="Add"/> <input type="button" value="Reset"/> <input type="button" value="Cancel"/>	

Edit Log

NSDUH Levels Lookup - Log

Variable Name	Year	After <input type="radio"/>			Modified Date	Before <input type="radio"/>		
<input type="text" value="drivalc2"/>	<input type="text" value="2004"/>	<input type="text" value="10"/>	<input type="text" value="30"/>	<input type="text" value="2005"/>	Between <input type="radio"/>	<input type="text" value="10"/>	<input type="text" value="31"/>	<input type="text" value="2005"/>
<input type="button" value="Filter"/>		<input type="button" value="Reset"/>		<input type="button" value="Main Page"/>		<input type="button" value="Logout"/>		

Search on Levels (5 Records)

Var Name	Study Year	Level Seq Number	Level	Definition	Ent By	Ent DT	Mod By	Mod DT	Log Type	Log DTTM
DRIVALC2	2004	3	4	Never/No Use (DRVALDR AND DRVAONLY=81,91,99)	Misty Foster	03/31/05	Misty Foster	04/18/05	update	04/18/05
DRIVALC2	2004	0	1	Yes (DRVALDR=1 OR DRVAONLY=1)	Misty Foster	03/31/05	Misty Foster	03/31/05	insert	03/31/05
DRIVALC2	2004	1	2	No (DRVALDR=2,89,99 AND DRVAONLY=2)	Misty Foster	03/31/05	Misty Foster	03/31/05	insert	03/31/05
DRIVALC2	2004	2	3	Unknown (OTHERWISE)	Misty Foster	03/31/05	Misty Foster	03/31/05	insert	03/31/05
DRIVALC2	2004	3	4	Never/No Use (DRVALDR AND DRVAONLY=81,89,91,99)	Misty Foster	03/31/05	Misty Foster	03/31/05	insert	03/31/05

Importing Documentation

Request Code Book / Import (Copy) Variables

Variable Name	Study	Year	Type	Section
<input type="text"/>	Main	2006 2005 2004	Select Final PUF	ALCOHOL BLUNTS COCAINE

For entire code book please send an email to Inga Allred or Susan Myers.

Please select the variables and click the button Request Code Book or Copy Variables to copy the data(labels,levels and comments) for the variables to another year. **Make sure the final list in the selected variables list is highlighted.**

8 records

Select From		Selected Variables		Code Book Copies to be Sent To: (type email ids separated by ;)	Special Note
BLNTEVER BLNTREC BLNT30DY BLNT30C1 BLNT30C2 RSNOMRJ RSNMRJMO BLNTNOMJ	>> << All > All <	<input type="text"/>	UP Dn	<input type="text"/>	<input type="text"/>
<input type="button" value="Copy Variables"/>		<input type="button" value="Request Code Book"/>		<input type="button" value="Reset"/>	

Ordering Variables

Order Variables			
Study	Year	Type	Section
Main	2000 2001 2002 2003 2004	Select Final Puf PUF Puf	BLUNTS SUBSTANCE DEPENDENCE AND ABUSE IMPUTED SUBSTANCE DEPENDENCE AND ABUSE RECODED SUBSTANCE DEPENDENCE AND ABUSE SPECIAL TOPICS
<input type="button" value="Filter"/> <input type="button" value="Reset"/> <input type="button" value="Main Page"/> <input type="button" value="Logout"/>			

Please Make sure the selected variables list is correctly ordered and highlighted.
8 records

Select From		Selected Variables	
1. BLNTEVER 2. BLNTREC 3. BLNT30DY 4. BLNT30C1 5. BLNT30C2 6. RSNOMRJ 7. RSNMRJMO 8. BLNTNOMJ	>> << All > All <		UP Dn
Make sure the final list in the selected Variables list is ordered and highlighted.			<input type="button" value="Save Order"/>

Documentation Search

NSDUH Variables Lookup

Variable Name	Study	Key Word Search in	Year	Type	Section	
<input type="text" value="alcever"/>	<input type="text" value="Choose one"/>	<input type="text" value=""/>	<input type="text" value="All"/> <input type="text" value="Labels"/> <input type="text" value="Comments"/>	<input type="text" value="Select"/> <input type="text" value="2005"/> <input type="text" value="2004"/>	<input type="text" value="Select"/> <input type="text" value="Final"/> <input type="text" value="PUF"/>	<input type="text" value="Select"/> <input type="text" value="ADOLESCENT DEPRESSION"/> <input type="text" value="ADULT DEPRESSION"/>
Sort Alphabetically <input type="checkbox"/>		<input type="button" value="Filter"/> <input type="button" value="Reset"/> <input type="button" value="Main Page"/> <input type="button" value="Logout"/>				

Search on Labels (1 Records - Code Book Order)

Var Name	Derived From	Category	Type	Length	Label	Low Range	Hi Range	Year	Raw Section	Analytic Section	PUF Section	Question (comment)	Footnote	Copy Variable	Master Delv
<u>ALCEVER</u>		1	N	2	EVER HAD DRINK OF ALCOHOLIC BEVERAGE			2004		ALCOHOL	ALCOHOL	<u>Comment</u>	None	<input type="button" value="Copy This"/>	<u>Delivered?</u>

Variable Delivery

NSDUH Variable Delivery Form

Year	2006 <input type="button" value="v"/>
Study	Choose one <input type="button" value="v"/>
Variables (separate by one space): <input type="button" value="Count"/> Please limit to 950 characters to fit in the email	<input type="text"/>
Variables that are NOT to be delivered to Client (separate by one space): <input type="button" value="Count"/> Please limit to 950 characters to fit in the email	<input type="text"/>
Redelivery?	<input type="checkbox"/> For Redeliveries please check this box
Date	03/10/2006
Person Delivered	Inga Allred
Data File Location	<input type="text"/>
Location of Frequencies	<input type="text"/>
Location of Map	<input type="text"/>
Notes <input type="button" value="Count"/> Please limit to 950 characters to fit in the email	<input type="text"/>
Schedule Line Number	<input type="text"/>
<input type="button" value="Submit"/> <input type="button" value="Reset"/> <input type="button" value="Main Page"/> <input type="button" value="Logout"/>	

Delivery History

NSDUH Variable Delivery Lookup

Variable Name	Study	Year	Person Delivered	After <input type="radio"/>	Modified Date	Before <input type="radio"/>		
txltystm	Choose one	2003		3 9 2006	Between <input type="radio"/>	3 10 2006		
New Vars? <input type="radio"/> Redelivered Vars? <input type="radio"/>			Delivered to Client <input type="checkbox"/>		Filter	Reset	Main Page	Logout

Search on Delivery History (3 Records)

Var Name	Study Year	Study	Date Delivered	Person Delivered	Data File Location	Delivered to Client	Freq Location	Map Location	Notes	Schedule Line Number	Redelivry
TXLTYSTM	2003	Main	02/24/04	Chuchun Chien	\\rtints29\nsduh9\edit\cai2003\sds\c03tmt14.sas7bdat	Yes	\\rtints29\nsduh9\edit\cai2003\delv\pfrq_cnttmt_14.lst	\\rtints29\nsduh9\edit\cai2003\macpgms\tmtmap.SAS	Yes		Yes
TXLTYSTM	2003	Main	02/18/04	Chuchun Chien	\\rtints29\nsduh9\edit\cai2003\sds\c03tmt14.sas7bdat	Yes	\\rtints29\nsduh9\edit\cai2003\delv\pfrq_cnttmt_14.lst	\\rtints29\nsduh9\edit\cai2003\macpgms\tmtmap.SAS	Yes	71	No
TXLTYSTM	2003	6monthMain	01/09/04	Chuchun Chien	\\rtints29\nsduh9\edit\cai2003\sds\c03tmt12.sas7bdat	Yes	\\rtints29\nsduh9\edit\cai2003\delv\pfrq_cnttmt_12.lst	\\rtints29\nsduh9\edit\cai2003\macpgms\tmtmap.SAS	Yes		No

Tracking (Re)delivery Reasons

Var Name	Note
TXLTYSTM	1. Documentation from 2002 has been moved into the tracking system for 2003.
TXLTYSTM	1. Please drop and pick these variables back up. These variables are being redelivered to change the data for 1 case who keyed all 1s in the remainder of his noncore ACASI questions, beginning with Specialty Cigarettes.

Cancel

Mapping of Variables

Child(ren) of Variable [TXLTILL2](#) for Year 2003

- 1: [ILLPCAIID](#) (Recoded) [View Details](#)
- 2: [ILLPCARE](#) (Recoded) [View Details](#)
- 3: [ILLPCORT](#) (Recoded) [View Details](#)
- 4: [ILLPEMPL](#) (Recoded) [View Details](#)
- 5: [ILLPFMLY](#) (Recoded) [View Details](#)
- 6: [ILLPINS](#) (Recoded) [View Details](#)
- 7: [ILLPMILC](#) (Recoded) [View Details](#)
- 8: [ILLPPUBP](#) (Recoded) [View Details](#)
- 9: [ILLPSAVE](#) (Recoded) [View Details](#)

Parent(s) of Variable [TXLTILL2](#) for Year 2003 (Type:Recoded)

- 1: [TXLTANL2](#) (Direct) [View Details](#)
- 2: [TXLTCOC2](#) (Direct) [View Details](#)
- 3: [TXLTHAL2](#) (Direct) [View Details](#)
- 4: [TXLTHER2](#) (Direct) [View Details](#)
- 5: [TXLTINH2](#) (Direct) [View Details](#)
- 6: [TXLTMJ2](#) (Direct) [View Details](#)
- 7: [TXLTSED2](#) (Direct) [View Details](#)
- 8: [TXLTSTM2](#) (Direct) [View Details](#)
- 9: [TXLTTRN2](#) (Direct) [View Details](#)

Requesting a Codebook

Request Code Book / Import (Copy) Variables

Variable Name	Study	Year	Type	Section
<input type="text"/>	Main	2006 2005 2004	Select Final PUF	ALCOHOL BLUNTS COCAINE
For entire code book please send an email to Inga Allred or Susan Myers. <input type="button" value="Filter"/> <input type="button" value="Reset"/> <input type="button" value="Main Page"/> <input type="button" value="Logout"/>				

Please select the variables and click the button Request Code Book or Copy Variables to copy the data(labels,levels and comments) for the variables to another year. **Make sure the final list in the selected variables list is highlighted.**

8 records

Select From		Selected Variables		Code Book Copies to be Sent To: (type email ids separated by ;)	Special Note
<input type="text"/>	>> <<	BLNTEVER BLNTREC BLNT30DY BLNT30C1 BLNT30C2 RSNOMRJ RSNMRJMO BLNTNOMJ	UP Dn	<input type="text"/>	<input type="text"/>
	All > All <				
<input type="button" value="Copy Variables"/>		<input type="button" value="Request Code Book"/>		<input type="button" value="Reset"/>	

Automatic Codebook Creation

- To ensure that all pertinent information is correct, the user requests a codebook online by choosing a set of variables.
- This request is processed on the server with a SAS program.
- The resulting codebook is emailed to the requestor within few minutes.
- Breaking down the review to smaller sections prior to the full review has greatly improved the quality and overall efficiency of the process.

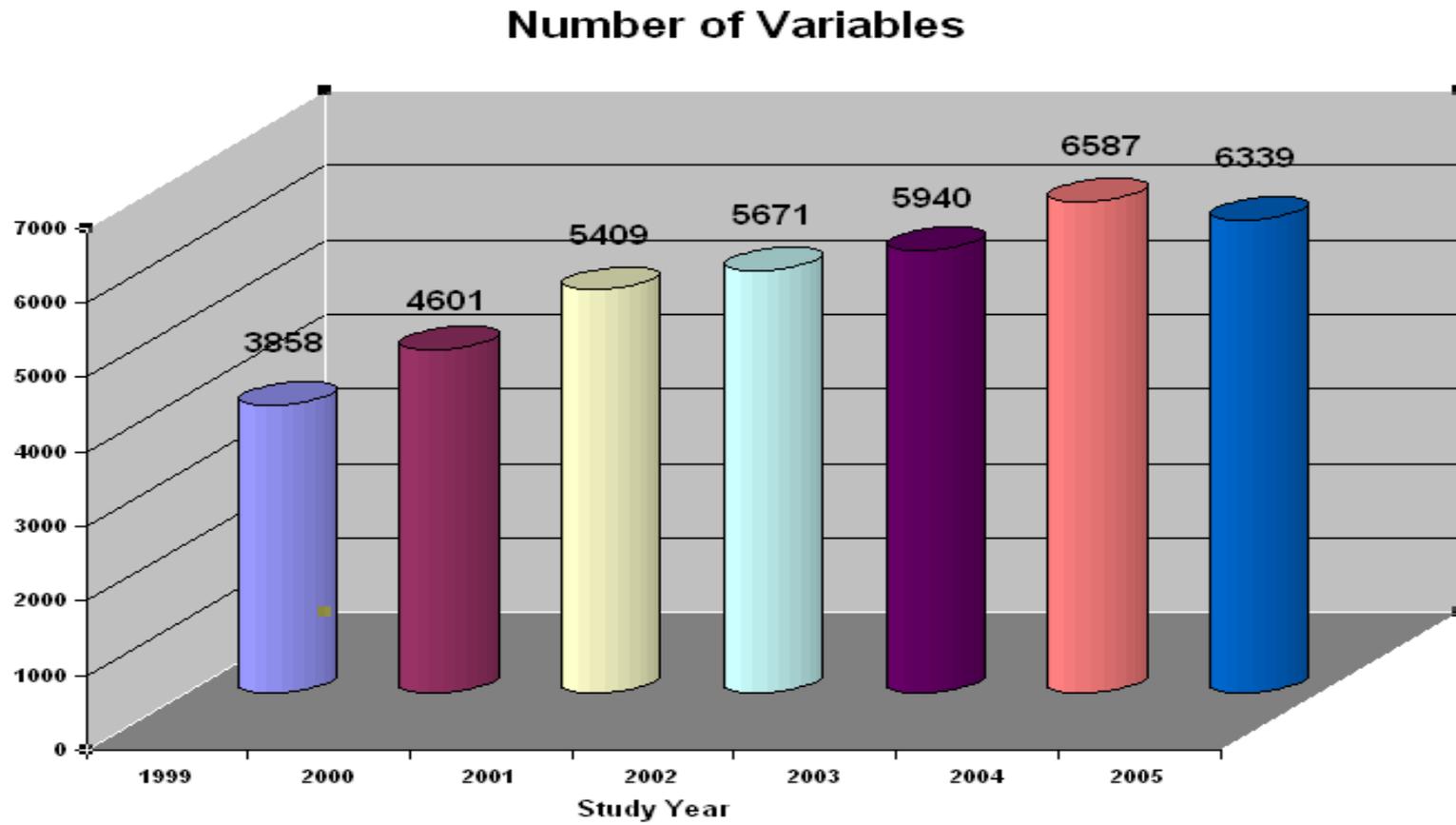
Benefits

- Multiple Users
- Quality Assurance
- Central Location available to everyone for use at any time
- Data Tracking
- Keyword searching
- Copy features
- Quicker response time to Client's requests

Rising Number of Variables

- Number of variables increase annually
 - Size of instrument increases
 - More analytic variables created
- VTS becomes more invaluable with the larger volume of variables

Number of Variables Per Year



Future Enhancements

- Implement a web based questionnaire development system. This would help in initial loading of question text to VTS.
- Create a utility for search and replace text within the descriptions of multiple variables across multiple years.
- Run and update frequencies automatically as and when the user requests them.
- Allow users to modify the Master Dataset through the system.
- Whatever else the users identify or wish that is valuable to the system.

Contact Information

Nanthini Ganapathi - ganapathi@rti.org	(919) 541-7143
Susan Myers - smyers@rti.org	(919) 541-7441
Inga Allred - irb@rti.org	(919) 541-7072

All authors can be reached at the mailing address:

RTI International

P.O. Box 12194

Research Triangle Park, NC 27709-2194

Client Disclaimer

- “The views expressed in written conference materials or publications and by speakers and moderators at HHS-sponsored conferences, do not necessarily reflect the official policies of the Department of Health and Human Services; nor does mention of trade names, commercial practices, or organizations imply endorsement by the U.S. Government.”